

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-305853

(43)Date of publication of application : 22.11.1996

(51)Int.Cl.

G06T 7/00

G06T 1/00

(21)Application number : 07-105318

(71)Applicant : MITSUBISHI ELECTRIC CORP

(22)Date of filing : 28.04.1995

(72)Inventor : KAGE YASUSHI
SHIONO SATORU
YAMADA SATOSHI

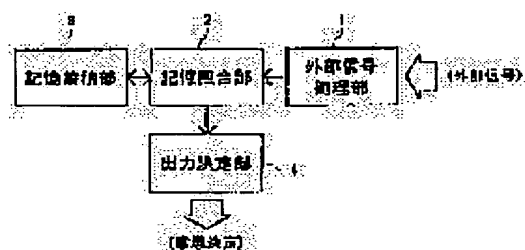
(54) METHOD AND DEVICE FOR OBJECT RECOGNITION AND DECISION MAKING BASED UPON RECOGNITION

(57)Abstract:

PURPOSE: To enable object recognition and decision making without any human intervention by automatically generating internal data representation that can be collated from an external signal.

CONSTITUTION: This device consists of four parts which are an external signal processing part 1, a storage matching part 2, a storage part 3, and an output decision part 4. This information processor extracts necessary information from the external signal by the external signal processing part 1 and converts it into internal data representation by attributes such as colors and shapes. By this internal data representation, an image being stored representation in the storage part 3 is retrieved and the object is estimated. Further, the output decision part 4 manages a correspondence table which is stored in the storage part 3 and has correspondence between the object and an action that should be taken to the object, and the action is determined on the basis of the correspondence table.

Thus, the object obtained as a two-dimensional shape image is estimated by using the image and an intention decision is flexibly made according to the state.



*** NOTICES ***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]An external signal treating part which incorporates sensor information about a new recognition object as an external signal, extracts attribution information of said recognition object from this external signal, and is changed into in-house-data expression which can compare this attribution information, A memory accumulating part which constitutes and memorizes in-house-data expression expressing conditional appearing probability of two or more attribution information related with each of two or more of said recognition objects, An information processor provided with a memory collating part which compares in-house-data expression of attribution information of a recognition object inputted from said external signal treating part and in-house-data expression of stored data of said memory accumulating part is used, An object identification method which presumes that a recognition object newly inputted from said external signal treating part among recognition objects memorized by said memory accumulating part and a recognition object which has a large in-house-data expression of similarity are said recognition object inputted newly.

[Claim 2]The object identification method according to claim 1 using an image specimen procession which consists of conditional appearing probability of each attribution information for every object as in-house-data expression of said memory accumulating part.

[Claim 3]The object identification method according to claim 1 or 2 which performs in study correction of conditional appearing probability which is stored data of said memory accumulating part, an addition of an attribution information item, or an addition of a recognition object based on a success and failure in said estimation processing according to experience of said estimation processing.

[Claim 4]As opposed to a recognition object which equips said information processor with a hypothetical reasoning part, in which said presumption is impossible and which was inputted newly, Perform abduction presumption using each in-house-data expression with other attribution information which was not used for said presumption, and stored data of said memory accumulating part, and candidates for presumption are enumerated, An object identification method given in any 1 paragraph of claims 1-3 which narrow down said candidate by similarity comparison of each in-house-data expression with said new input-pairs elephant and said each candidate object.

[Claim 5]Inside of a recognition object characterized by comprising the following memorized by said memory accumulating part using an information processor, From said external signal treating part, presume that a recognition object inputted newly and a recognition object which has a large in-house-data expression of similarity are said recognition object inputted newly, and it is recognized, A decision-making method which opts for action which should be made to said recognized recognition object from a conversion table of a recognition object of said output deciding part, and action which should be made, and outputs this as a command.

An external signal treating part which incorporates sensor information about a new recognition object as an external signal, extracts attribution information of said recognition object from this external signal, and is changed into in-house-data expression which can compare this attribution information.

A memory accumulating part which constitutes and memorizes in-house-data expression expressing conditional appearing probability of two or more attribution information related with

each of two or more of said recognition objects.

A memory collating part which compares in-house-data expression of attribution information of a recognition object inputted from said external signal treating part, and in-house-data expression of stored data of said memory accumulating part.

An output deciding part which has a conversion table with action which should be made to a recognition object and each recognition object of said plurality, and outputs as a command said action which should be made.

[Claim 6] Equip said information processor with an output evaluating part, and the validity of said command or said recognition is evaluated from a reaction of a recognition object of a result by which said command was executed, A decision-making method according to claim 5 of performing in study correction of conditional appearing probability which is stored data of said memory accumulating part, an addition of an attribute item, or an addition of a recognition object based on this evaluation.

[Claim 7] A decision-making method according to claim 5 or 6 restricted to the contents for emergencies which set up beforehand the contents of said estimation processing and said decision-making processing based on accident information from environment where equip said information processor with a condition-monitoring part, and said decision-making is performed, or accident information from said sensor.

[Claim 8] An object recognition device which presumes that a recognition object is said recognition object inputted newly, comprising:

An external signal treating part which incorporates sensor information about a new recognition object as an external signal, extracts attribution information of said recognition object from this external signal, and is changed into in-house-data expression which can compare this attribution information.

A memory accumulating part which constitutes and memorizes in-house-data expression expressing conditional appearing probability of two or more attribution information related with each of two or more of said recognition objects.

It has a memory collating part which compares in-house-data expression of attribution information of a recognition object inputted from said external signal treating part, and in-house-data expression of stored data of said memory accumulating part, Large in-house-data expression of a recognition object and similarity newly inputted from said external signal treating part among recognition objects memorized by said memory accumulating part.

[Claim 9] The object recognition device according to claim 8 using an image specimen procession which consists of conditional appearing probability of each attribution information for every object as in-house-data expression of said memory accumulating part.

[Claim 10] The object recognition device according to claim 8 or 9 which ** is in experience of said estimation processing, and performs in study correction of conditional appearing probability which is stored data of said memory accumulating part, an addition of an attribution information item, or an addition of a recognition object based on a success and failure in said estimation processing.

[Claim 11] It has a hypothetical reasoning part which performs abduction presumption using each in-house-data expression with other attribution information which was not used for said presumption, and stored data of said memory accumulating part to a recognition object in which said presumption is impossible, and which was inputted newly, and enumerates candidates for presumption, An object recognition device given in any 1 paragraph of claims 8-10 which narrow down said candidate by similarity comparison of each in-house-data expression with said new input-pairs elephant and said each candidate object.

[Claim 12] A decision-making device which presumes that a recognition object characterized by comprising the following is said recognition object inputted newly, recognizes it, opts for action which should be made to said recognized recognition object from a conversion table of a recognition object of said output deciding part, and action which should be made, and outputs this as a command.

External signal processing which incorporates sensor information about a new recognition object as an external signal, extracts attribution information of said recognition object from this

external signal, and is changed into in-house-data expression which can compare this attribution information.

A memory accumulating part which constitutes and memorizes in-house-data expression expressing conditional appearing probability of two or more attribution information related with each of two or more of said recognition objects.

A memory collating part which compares in-house-data expression of attribution information of a recognition object inputted from said external signal treating part, and in-house-data expression of stored data of said memory accumulating part.

Large in-house-data expression of a recognition object and similarity newly inputted from said external signal treating part among recognition objects which have a conversion table with action which should be made to a recognition object and a recognition object of said plurality, were provided with an output deciding part which outputs as a command said action which should be made, and were memorized by said memory accumulating part.

[Claim 13]It has an output evaluating part which evaluates the validity of said command or said recognition from a reaction of a recognition object of a result by which said command was executed, The decision-making device according to claim 12 which performs in study correction of conditional appearing probability which is stored data of said memory accumulating part, an addition of an attribute item, or an addition of a recognition object based on this evaluation.

[Claim 14]The decision-making device according to claim 12 or 13 provided with a condition-monitoring part restricted to the contents for emergencies which set up beforehand the contents of said estimation processing and said decision-making processing based on accident information from environment where said decision-making is performed, or accident information from said sensor.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application]This invention extracts required information from an external signal, by performing collation with memory, association, and reasoning about that information, recognizes an object according to a situation and relates to the method of the information processing which makes decisions, and its device.

[0002]

[Description of the Prior Art]Many researches for making a computer execute by proxy decision-making which human being performs using the information from the external world have so far been made. They are roughly divided into the artificial intelligence research based on sign reasoning, and image engineering research aiming at object recognition etc. Although the external signal acquired from the external world is not limited, its rate of being used is dramatically large, and it is limited and described to a picture signal as a picture signal here. Many inference systems, such as an expert system, were built and it has been put in practical use as a decision support system by the former artificial intelligence research. The latter image engineering research has established the area of research called a computer vision.

[0003]There is a knowledge base inference system which provides with and has both a rule base by Yasunobu 1000 **** and others published by Japanese Society for Artificial Intelligence (1992), Vol.7, No.6, and pp1087-1095 and a case base as an example of an inference system. Many of conventional expert systems store in a rule base the knowledge which a specialist has, and they reason based on the rule. However, generally, on the examination business of a business field, it generates mostly, and in that case, the situation which is not provided in examination regulation also refers to the past examination example, and is judged. Thus, it is useful for judgment to combine with a rule and to utilize an example in many cases, and the system of this conventional example refers to a rule or an example according to a situation, and supports decision-making.

[0004]The composition of this system is shown in drawing 7. The rule base which has memorized the knowledge in which a specialist has 8 as an if-then rule, and 9 are case bases which have memorized the problem and the solution to past as an example. 10 is a working memory and has memorized the data of the present problem, a solution, a middle result, etc. 11 is an inference engine and has the following three functions.

[1]Selection of a reasoning paradigm: About any of the reasoning based on the rule base 8 or the case base 9 are chosen, collation processing with the contents of the working memory 10 determines, and perform suitable reasoning to the problem faced now. The continuous value between 0 and 1 is computed as goodness of fit for whether the example which constitutes the rule which constitutes the rule base 8, or the case base 9, and the contents of the working memory 10 suit. Based on this collated result, reasoning based on each is performed by whether it is a rule of the rule base 8 which has the highest goodness of fit, or it is an example of the case base 9. Priority is given to the reasoning based on the rule base 8 when both rules are the same goodness of fit.

[2]Search service: Search the example which searches a suitable if-then rule from the rule base 8 to the present problem, or is similar to the present problem from the case base 9. Search is calculated as a continuous value between 0-1 about whether each rule or an example is useful

like the above-mentioned goodness-of-fit calculation.

[3]Execution function: Reason the solution of the present problem by processing the example which was performed according to the if-then rule searched from the rule base 8, or was searched from the case base 9 so that the present problem may be suited.

[0005]12 is a user interface and is a device for a user to perform series, such as setting out and change of the data of the working memory 10, directions of starting of reasoning, and a check of an inference result.

[0006]The above-mentioned knowledge based system was begun, in many inference systems, it could reason only within the limit of deductive inference, but the extendibility of a reasoning system had a problem. In order to solve this, various reasoning gestalten have been advocated from theoretical research of sign reasoning. The reasoning gestalt especially of abduction attracts attention in recent years, and research is advanced.

[0007]Generally, when placed by the strange situation, human being builds up a hypothesis based on the knowledge acquired by the past experience, and makes a suitable judgment according to a situation. Abduction is based on the reasoning process based on such a hypothesis. It is possible to newly make a hypothesis from empirical knowledge, to verify the validity by trial and error unlike inductive inference deductive [conventional], by this, and to incorporate as new knowledge. Generally the form of abduction reasoning is given as follows. (1) $A \rightarrow B$ — and — If it becomes (2) $C \rightarrow B$ (3) $A \rightarrow C$ (1) shows the situation faced now, and corresponds to one in the knowledge which experienced (2) until now and is known, and (3) is the hypothesis drawn based on (1) and (2).

[0008]Since many C in (2) generally exists as a candidate who draws B, it becomes important from a point of processing speed for a number huge as a result of hypotheses to be acquired, therefore to narrow down the number. It is necessary to inspect whether it is contradictory to the logical system which the hypothesis furthermore acquired already has. Based on two of the hugeness of the number of such hypotheses, and the inspection of consistency, there is an example by Katsumi Inoue and others who were published by Japanese Society for Artificial Intelligence (1993), Vol.8, No.6, and pp786-796 as a system which aimed at processing the abduction reasoning by symbolic logic at high speed.

[0009]Drawing 8 is the configuration of module of an abduction inference system. The translation module which changes into the sign expression for high speed calculation the logical formula with which 13 was described by predicate calculus, It is a consistency inspection module which inspects whether the hypothetical generating module which generates a hypothesis, and 15 are contradictory to the logical system to which the generated hypothesis is already given from sign expression from which 14 was changed. There are the following two kinds in the detection method of the inconsistency to a certain hypothesis.

(I) Reason based on the reasoning system described by the logical formula, and detect inconsistency.

(II) Register the hypothetical set which causes inconsistency into the reasoning system, and collation with it detects inconsistency. In this example, the consistency inspection based on above (II) is processed with the module 15, And abduction reasoning is made to speed up by increasing the number of a computer which is realized on the high-speed parallel computer for logical languages, and performs the function of the module 14 and the module 15 in parallel.

[0010]Each of two above-mentioned conventional examples is sign inference systems, and the following problems produce them. Since all information required for problem solving is given as sign expression, human being needs to perform all of the conversion in a sign from external world information, and grant of a meaning. Various sensor information, such as vision information, is processed, and if the method of expressing as knowledge is not given, any systems always need an intervention of human being. Therefore, it is impossible to make a computer execute by proxy all decision-making that human being performs, and a solvable problem is limited extremely.

[0011]Construction of the system which processes sensor information and changes external world information into an internal expression from this is indispensable. In image engineering research of a computer vision etc., objective shape and a motion were discriminated from a still picture or video, the model for separating a target from a still more complicated background was

advocated, and many image recognition systems have been built. There is an example by Junji Yamato and others to whom it was shown at JP,5-46583,A among these as an image recognition system which performs action recognition of human being by time varying image processing. With many image recognition systems, it aims at an advanced cognitive level called action recognition of human being by this conventional technology to having stopped at low cognitive levels, such as shape extraction.

[0012]Drawing 9 expresses the algorithm of this conventional example. First, the video which contains human being under action from the image input part 16 is caught, and it stores in the memory 17 for pictures. Next, a feature vector is obtained from video by the feature extraction part 18. An example of the calculation method of a concrete feature vector is explained here.

[0013]As shown in drawing 10, the memory for pictures is divided into the subblock of $N \times M$ with the number of pictures of $n \times m$, and binarization of a picture is performed by each subblock. Next, it asks for the share of the black pixel in this subblock, and is considered as the feature vector of the $N \times M$ dimension which makes the share a variable value. This feature vector is recorded on the feature storing memory 19. And a feature vector is changed into a symbol by the quantizing part 20, and is recorded on the symbol storing memory 21. This feature vector explains the method changed into a symbol.

[0014]Drawing 11 is a typical example of a picture for human being's operation recognition, and drawing 12 is a symbol (number) corresponding to the feature vector of each picture of drawing 11. After two or more image data incorporated into the time series system is changed into a feature vector, respectively, it is changed into the sequence of the symbol corresponding to a feature vector. Drawing 13 indicates the example from which the sequence of another image data was changed into symbol trains to be a certain drawing 11 incorporated serially. The quantizing part 20 has some representative point vectors beforehand, and calls these each a symbol. That is, with conversion as the symbol in the quantizing part 20, a feature vector is equivalent to choosing the symbol corresponding to a representative point vector with the nearest distance in a representative point vector group.

[0015]Next, from each of the model stored in the state transition model storing memory 25 for recognition for which only the number of categories to recognize was prepared, the probability that a feature vector will be generated is computed by the likelihood calculation part 22. The model in which the called-for likelihood serves as the maximum is chosen as a recognition result, and is accumulated in the memory 23 for recognition results. The model parameter estimating part 24 presumes the parameter of a state transition model which generates the symbol to the symbol obtained from the data for study given for every category, and accumulates it in the state transition model storing memory 25 for recognition. [two or more]

[0016]In this conventional example, four operations (the right hand which raises a left leg, which raises a left hand and which raises a right leg is raised) are mentioned as an example as action of a recognition object, and the effect of an average of about 90% is acquired as a recognition rate. However, the technique of operation of human being who operation of identifiable human being is limited to several kinds, and newly registers strange operation as one operation being limited to several kinds, and newly registering strange operation as one operation was not given, but there was a problem in the extendibility of the operation which can be recognized.

[0017]

[Problem(s) to be Solved by the Invention]The method of Yasunobu and others who is the examples of the conventional inference system is a deductive inference system based on an if-then rule, and there was a problem in the extendibility of a reasoning system. The method of Inoue and others who introduced abduction reasoning in order to solve the problem is coping with it by processing calculation of a huge quantity about hypothetical processing at high speed by an exclusive parallel computer. However, the problem of a hypothetical number of hugeness is not solved in essence. Since each of these two inference systems is furthermore sign inference systems, The problem that all of the conversion to the internal expression of the external world information caught by the problem, i.e., a various sensor, that human being has to give all information required for problem solving as sign expression, and grant of a meaning need an intervention of human being is left behind.

[0018]The method of Daiwa and others which is the examples of the conventional image

recognition system aims at a way method of recognition human being does not involve. However, operation of identifiable human being was limited to several kinds, and the correspondence to strange operation was not taken into consideration, but there was a problem in the pliability of operation recognition.

[0019] This invention is what was made in order that the image recognition system of the above conventional inference systems might cancel the problem held, respectively, An object is recognized without responding to the given situation flexibly and human being involving, and it aims at providing the information processing method which makes decisions based on the recognition, and its device.

[0020]

[Means for Solving the Problem] In order to solve the above-mentioned problem, it needed to be considered in detail how human being would make decisions by understanding the external world. Therefore, it came to propose a method of information processing doing research which took not only the results of research, such as artificial intelligence and image engineering, but knowledge of a field of physiology, psychology, and logic into consideration, and using memory expression (in-house-data expression) called pulsus cordis mentioned later, and its device.

[0021] An information processor concerning this invention makes pulsus cordis using information which stored in a storage parts store sensor information, such as vision information which contains a recognition object, for example, after extracting information required as an external signal, presumes a recognition object using that pulsus cordis, and performs decision-making according to a situation based on that result. An object estimation means of this information processor comprises at least three portions, an external signal treating part, a memory accumulating part, and a memory collating part.

[0022] An external signal treating part is stated in the first place. For example, when an external signal is vision information (it limits and states to vision information below since an external signal of a sound or other type, like it is stinking can be treated similarly), and when making in-house-data expression, a vision model based on physiological knowledge is adopted as a means to extract required information. It is possible by acquiring many knowledge from vision physiological research in the ape cerebral cortex in recent years, and copying the mechanism to realize a visual system near human being.

[0023] A vision related field of an ape is divided into some fields, and making a layered structure with the locality of analysis of retina information is known. Retina information, including shape, a motion, a color, etc., is analyzed locally, is gradually unified by other ****, and makes information for recognition judgment from primary visual area. So, in this invention, attributes, such as shape, a motion, a color, and a texture, are provided, a picture signal acquired with a CCD camera etc. is processed, information corresponding to an attribute is extracted, and it changes into in-house-data expression which can be compared within a system.

[0024] A memory accumulating part is stated to the second. In-house-data expression in a memory accumulating part uses a thing near human being's memory expression proposed by psychology research. When making a certain three dimensional object into a recognition object, a different shape image is acquired from a different viewpoint. Since a recognition object is generally observed from a specific viewpoint in many cases, a difference of probability arises in how whose object is visible. As a result, conditional probability has followed on each shape image which an object has.

[0025] In this invention, the object S reflecting vision experience of a different set a shape image's by difference in the way of being visible which the one object S has, and an observer's past uses as in-house-data expression [in / it doubles, is called pulsus cordis of the object S, and / a memory accumulating part] with conditional probability which is visible to a certain shape. Attribution information, such as a shape image which constitutes pulsus cordis, is built by in-house-data expression extracted by an external signal treating part.

[0026] It is equivalent to searching a shape image which constitutes pulsus cordis with search of a memory content of a memory accumulating part by in-house-data expression [in / a memory accumulating part] about shape. It is the same about other attributes, such as a motion. In-house-data expression which is a memory content of a memory accumulating part, and its appearing probability are not what was fixed, and are updated as required by experience.

[0027]A memory collating part is explained to the third. A candidate of a recognition object in a memory accumulating part is searched and extracted using similarity of attribution information which was made from an external signal treating part about a recognition object, and was changed into in-house-data expression, and attribution information of a memory accumulating part, and those sets are made to this memory collating part. It is the main roles of this memory collating part to perform operation which searches this set and makes a match in-house-data expression with a recognition object best.

[0028]According to research of cognitive psychology, memory of human being is roughly divided into two, short-term memory and long-term memory. Short-term memory is memory about an occurrence which happened these days, and long-term memory is almost fixed and is memory held through a lifetime. After a phenomenon laid on heart first is incorporated into short-term memory and strengthened by frequent remembrance as these roles, it is thought that it is fixed to long-term memory. A memory collating part and a memory accumulating part of this invention correspond to above-mentioned short-term memory and long-term memory, respectively.

[0029]Next, an output deciding part used for a decision-making means of this invention is described. An output deciding part opts for action which should be made to what was presumed to be a recognition object by a memory collating part. For example, if a recognition object is "fire", "fire fighting is performed" is matched, in addition if it is "a car which approaches", various behavior patterns are matched by recognition object, such as "avoiding." According to research of a cerebrophysiology, by human being's cerebral cortex, an object is recognized by a sensory association area, the frontal lobe receives the information, and it opts for action which should be made according to an object. An output deciding part in this invention is made to correspond to a function of this frontal lobe.

[0030]Generally action for which an output deciding part opts here corresponds with a known recognition object. However, it is necessary to make decisions by building up a hypothesis also to a case where information from an object is insufficient and recognition is difficult, or a strange object. It is necessary to evaluate whether action based on the decision-making was still more appropriate. In addition to the four above-mentioned sorts of main formation parts, in this invention, a hypothetical reasoning part and an output evaluating part are provided, when in-house-data expression is insufficient in the former and it cannot recognize, abduction reasoning is performed, a candidate of a recognition object is selected out of this, and the validity of action based on it is evaluated in the latter. It has a role which extends a recognition system which a memory accumulating part has by changing a memory content of a memory accumulating part concerning a known recognition object in an output evaluating part furthermore, or newly registering a strange object into a memory accumulating part.

[0031]In order for an accident to arise in an internal state of the target system and to recover it, by the usual information processing routine, it may become difficult to cope with it promptly. As a function corresponding to this, when a physiological phenomenon of dryness of a throat or everything in the living body is faced, a part called hypothalamus detects the change and being concerned with decision-making for shifting to prompt action is known for a living thing.

[0032]Emotional behavior based on information on hypothalamus copes with a case in an emergency in conformity to judgment which went via the cerebral cortex requiring time not passing through the cerebral cortex. A condition-monitoring part which supervises an accident of a parameter inside an object system or a situation of the object system exterior in this invention in addition to the main 4 above-mentioned formation parts is provided, An external signal treating part and a memory collating part are ordered to perform a change to an emergency manipulation routine, and it makes it possible to cope with an accident of a system internal state in conformity.

[0033]

[Function]Since the information processor by this invention was constituted from an external signal using the model imitating the information processing function which a living thing has as a technique for changing into in-house-data expression, it can gain in-house-data expression, without human being intervening. Since the expressive form has different redundancy from the conventional sign expression, reasoning more flexible than sign reasoning is possible for it, and it

can solve the problem of the computational complexity which was seen by the abduction reasoning by symbolic logic and which increases explosively.

[0034] Since it constituted so that judgment over a strange recognition object might be enabled, according to a situation, extension and correction are added to a memory system, and decision-making based on flexible object recognition and it is performed. It has the mechanism for this information processor to cope with the critical situation where the target system may be faced, in conformity furthermore. It not only can use it as an independent decision-making device, but it can use this information processor for the common system which needs decision-making, building it into it as a decision-making device.

[0035]

[Example]

Example 1. drawing 1 is an example of the device in which the composition of this example is shown, and consists of four portions of the external signal treating part 1, the memory collating part 2, the memory accumulating part 3, and the output deciding part 4. In the information processor by this invention, required information is extracted from an external signal by the external signal treating part 1, and it is changed into in-house-data expression for every attributes, such as a color and shape. And with this in-house-data expression, the pulsus cordis which is memory expression of the already described memory accumulating part 3 is searched, and an object is presumed.

[0036] human being — after the birth — various two-dimensional shape images which a certain specific object has are memorized with the appearing probability through various vision experiences, and a recognition judgment is made on recognition of a subsequent object in consideration of the appearing probability based on vision experience of a shape image and the past. Thus, what expressed the correspondence relation between some objects and various shape according to appearing probability is called an image specimen procession (Science (1992), Vol.257, pp1357–1363). When three dimensional object S_1 , —, conditional probability that object S_1 can have seen S_n as retinal image L_j when being reflected as L_m , retinal image L_1 , —, are set to $P(L_j|S_i)$, an image specimen procession is expressed like drawing 2.

[0037] Human being object S_1 , —, shape image L_1 to S_n , —, by gaining correlation of L_m as a procession of conditional probability like drawing 2, and referring to this procession to a line writing direction in the case of the visual perception after acquisition, From specific shape image L_j , the original object presumes [S_1 , —,] any of S_n they are. The probability $P(S_i|L_j)$ that the original object when shape image L_j is observed is S_i is given with a following formula by Bayesian estimation using an image specimen procession.

$$P(S_i|L_j) = P(L_j|S_i) P(S_i) / [P(L_j|S_1) P(S_1) + \dots + P(L_j|S_n) P(S_n)] \quad (1)$$

Here, $P(S_1)$ and $\dots P(S_n)$ show the occurrence probability of object S_1 , —, S_n , respectively, and are expressed with a following formula.

$$P(S_i) = P(L_1|S_i) + \dots + P(L_m|S_i) \quad (2)$$

the value of each probability is reflecting the past vision experience — a situation — an anacritic value is taken.

[0038] After being based on the above view, the presumed method of the three dimensional object in this example is considered as follows. Three dimensional object S_x has a shape set $\{L_1, \dots, L_y\}$ of the two-dimensional image from which how for it to be visible by the difference among conditions, such as a direction to see, differs. Each shape image L_j is connected by the conditional probability P reflecting vision experience of object S_x and the past ($L_1|S_x$), —, $P(L_y|S_x)$. This shape set is called the pulsus cordis of object S_x . When presuming the original object from observed shape image L_j , all the objects that contain shape L_j in the pulsus cordis are made into a candidate, and the presumed probability $P(S_i|L_j)$ is calculated from (1) type about all candidate-objects S_i ($i=1, \dots, n$). The object S in which the probability serves as the maximum is a three dimensional object presumed from L_j .

[0039] This example explains object body presumption based on a shape image, and the action determination to it. That is, priority is given to presumption by shape among various attributes which vision information has, and other attributes, such as a motion and a color, presuppose that a priority is low in the case of object presumption.

[0040] The external signal treating part 1 comprises a CCD camera and an image processing device. The picture photoed by the former is processed using the latter, and the shape L is extracted from the two-dimensional image of recognition objects.

[0041] The memory accumulating part 3 has remembered the appearing probability $P(L_j|S_i)$ to be shape image L_j ($j=1, \dots, m$) obtained using the external signal treating part 1 of object S_i ($i=1, \dots, n$) made into a recognition object as in-house-data expression. That is, when L_j is beforehand obtained using the external signal treating part 1 and the appearing probability also asks for various shape of S_i automatically under the environment which actually applies this example, the memory of the memory accumulating part 3 is built. The memory content of this memory accumulating part has incorporated a learning mechanism which is changed in the form of the addition of a shape image, or change of appearing probability by future experiences. Change of appearing probability increases, when it succeeds, for example in the target presumption, and it is made to decrease when it fails. Judgment of failure and a success is performed by whether the reference value of the similarity set up beforehand is exceeded.

[0042] When not succeeding in the target presumption, the object is registered as a strange object, and appearing probability is built by subsequent presumed experience. Thus, the object to presume is extensible in study.

[0043] The memory collating part 2 searches the memory accumulating part 3 to L obtained by the external signal treating part 1, and enumerates all object S_i ($i=1, \dots, n$) included in a set by making L more than fixed probability into shape. From the appearing probability $P(L|S_i)$

furthermore memorized by the memory accumulating part 3, the candidate S of the original object is presumed according to (1) type.

[0044] In the output deciding part 4, the action taken to each object is managed as a conversion table. The action A which should be taken to the object S of the origin presumed from the shape L by the memory collating part 2 is chosen by the output deciding part 4.

[0045] An animal is supervised in the animal park which has bred the primates for concrete operation of these formation parts specially, warning is emitted to the animal which is present in a dangerous place, and the example of the dangerous place managerial system carrying the system made to leave the spot explains. If the animal observed is human being, cautions will be urged to it in human being's voice, this system will tell a shooting sound, if it is a gorilla, and if it is a chimpanzee, it is a system for sounding and driving off an alarm whistle. Each animal other than human being assumes that a condition is attached to corresponding sonic stimulation so that the spot may be made to leave.

[0046] The external signal treating part 1 is what equipped the CCD camera and the image information processing unit for shape extraction, records the conditional probability that the shape of various animals seen in the past will be observed, and, on the other hand, is making the memory accumulating part 3 memorize the result.

[0047] The external signal treating part 1 extracts the shape L of **** and an object for an object first. The memory collating part 2 searches memory of the memory accumulating part 3 based on the extracted shape L, and enumerates as a candidate what is more than the value which the probability $P(S)$ which exists in the zoo of the object S set up about the object which has the shape L in the pulsus cordis.

[0048] In this example, the shape L enables Lknown shape L_1, \dots, L_m , and specification. The shape L is compared [that it can specify and] with all known shape, as a result of calculating similarity with each shape, it becomes beyond the reference value which the value of one of similarity set up, and it calls the shape which gives the maximum of similarity in it the shape specified from the shape L.

[0049] As a result of analyzing an object, supposing human being, a chimpanzee, and gorillas are enumerated as a candidate, the original object will be presumed by the probability and (1) type

which are memorized by the memory accumulating part 3. The probability $P(L|human)$, $P(L|chimp)$, and $P(L|goril)$ that human being, a chimpanzee, and a gorilla will be extracted as the shape image L , respectively. The probability $P(human|L)$ which can be presumed that the original object is human being from the shape L with the probability $P(human)$, $P(chimp)$, and $P(goril)$ that human being, a chimpanzee, and a gorilla exist in a zoo, respectively is calculated by a following formula.

$$P(human|L) = \frac{P(L|human) \cdot P(human)}{P(L|human) \cdot P(human) + P(L|chimp) \cdot P(chimp) + P(L|goril) \cdot P(goril)}$$

The same may be said of the probability $P(chimp|L)$ and $P(goril|L)$. As a result, if the value of $P(human|L)$ is the maximum, for example, it will be presumed that the original object is human being and this information will be transmitted to the output deciding part 4.

[0050] In the output deciding part 4, the memory accumulating part 3 memorizes, the matching table of an object and the action which should be taken to it is managed, and action determination is made based on this. For example, the output deciding part 4 assumes that a matching table like drawing 3 is managed. The action "the warning by human being's voice" is chosen from this table.

[0051] Thus, it is possible to presume the object regarded as a two-dimensional shape image using the pulsus cordis, and to give flexible decision-making according to a situation.

[0052] The composition of the example 2. example 2 is the same as drawing 1. In Example 1, when shape image L_j was observed, the probability $P(S_i|L_j)$ which can be presumed that the original object is S_i was calculating by the Bayesian estimation given by (1) formula. and object S_1 , —, the occurrence probability S of $S_n(S_1)$, —, the situation where $P(S_n)$ reflected the past vision experience — it was an anacletic value. In this example, it is calculating noting that all the appearing probability of this each set elephant is the same values. As a result, (1) type is simplified like a following formula.

$$P(S_i|L_j) = \frac{P(L_j|S_i)}{P(L_j|S_1) + \dots + P(L_j|S_n)} \quad (3)$$

[0053] The presuming method by this example has calculation quicker than presumption by (1) type, and each set elephant S_i ($i = 1, \dots, n$) is the effective presuming method in the situation of appearing almost similarly etc. Each set elephant will be compared impartially.

[0054] The composition of the example 3. example 3 is the same as drawing 1. Example 1 showed the example which presumes the recognition object S by the shape image L . However, it is possible to use together other attribution information other than shape, such as the color C , for example, and to raise the reliability over presumption of the recognition object S by presumption by a shape image, in the picture captured with the CCD camera etc., since it is unreliable when there are many noises. At this example, by the external signal treating part 1, it has an image processing portion which considers detection of a color as shape, and the memory accumulating part 3 presupposes that shape and a color are memorized with the conditional probability reflecting appearing probability about each recognition object.

[0055] The same example of application as Example 1 explains. The shape L and the color C are extracted by the external signal treating part 1, and the memory collating part 2 detects memory of the memory accumulating part 3. By object presumption by shape, human being, a chimpanzee, a gorilla, and a gibbon presuppose that human being, a chimpanzee, a gorilla, and Japanese macaques were enumerated as candidate management by object presumption by a color. Then, suppose that presumed probability was calculated about the shape L and the color C , and the following results were obtained about human being, chimpanzee, and gorilla which are three common candidates.

Shape L : $P(human|L) = 0.20$, $P(chimp|L) = 0.30$, $P(goril|L) = 0.20$ color C : $P(human|C) = 0.70$, $P(chimp|C) = 0.10$, $P(goril|C) = 0.10$ [0056] In this case, although it may judge that the recognition object S is a "chimpanzee" only using the presumed probability of the shape L like Example 1, when the reliability of object presumption by shape is low, the presumed probability of the color C is also calculated by using it simultaneously. For example, the average of the shape L and the presumed probability of the color C is calculated, and, in the case of "human being", in the case of : $\{P(human|L) + P(human|C)\} / 2 = 0.45$ a "chimpanzee", it is : $\{P(chimp|L)$

+P (chimp |C)) / 2=0.

case [of 20 a "gorilla"]: — {P(goril |L) +P (goril |C)) / 2=0.15 — it is possible to presume that an object is "human being" by this calculation result.

[0057]This example showed the example of presumption which used the information on shape and a color. However, when a luminosity is not enough, it is also possible judgment by a color and to raise the reliability of object presumption using the information on a motion, since it is unreliable in many cases. In this example computation, although the average of the probability about each attribute is calculated as an arithmetical average, it is also possible to evaluate shape, and a color and the reliability to a motion according to a situation, and to calculate as a weighting average.

[0058]Thus, the reliability of object presumption in the memory collating part 3 can be greatly raised by using the pulsus cordis of attributes, such as not only shape but a motion, a color, etc.

[0059]Example 4. drawing 4 is an example of the device in the case of having the output evaluating part 5 also as a formation part in addition to the composition which showed the equipment configuration of this example and was shown in Example 1. The output evaluating part 5 is provided with the function which supervises whether decision-making given by the output deciding part 4 is the right.

[0060]The case where the action "the warning by human being's voice" is chosen by the output deciding part 4 is considered to the object S observed in the same situation as Example 1. The output evaluating part 5 observes action of the object S to this warning. Since a condition is attached so that it may leave by respectively individual sonic stimulation if it is animals other than human being, an animal cannot be made to react to human being's voice immediately, and to leave the spot promptly, if the object S is human being, but to leave in human being's voice.

[0061]If it is judged by the output evaluating part 5 that there is no change in action of the object S, the output evaluating part 5 will decrease the value of the probability P (last shipment) that the object S will be observed by the external signal treating part 1 as the shape image L in the image specimen procession memorized by the memory accumulating part 3. Thus, processing is urged to the output evaluating part 5 to the memory collating part 2 so that the object S may be presumed once again based on the updated image specimen procession. Even if it updates an image specimen procession, when not succeeding in the target presumption, the object is registered as a strange object and appearing probability is built by subsequent presumed experience. Although renewal of an image specimen procession and extension for new were started by a presumed success and failure by a similarity standard in Example 1, Since it can start based on the action identification result of the object by the output evaluating part 5 in this example, it becomes correction of a more reliable image specimen procession extensible.

[0062]Thus, by judging whether it is the action the action of the object S based on decision-making given by the output deciding part 4 was predicted to be by the output evaluating part 5 by the output evaluating part 5, It becomes possible to realize the learning function which corrects the memory system of the memory accumulating part 3, is flexibly equivalent to the given situation, and improves presumed reliability.

[0063]Example 5. drawing 5 is an example of the device in the case of having the hypothetical reasoning part 6 as a formation part in addition to the composition which showed the equipment configuration of this example and was shown in Example 1. When the target recognition is impossible for the function of the hypothetical reasoning part 6, it is building up a hypothesis about what it being. Although this example presumes an object like Example 3 based on the shape and the color which were detected by the external signal treating part 1, Differing from Example 3 is the point of assuming the case where it is able to be specified as neither of the two-dimensional shape as a result of the shape L extracted by the external signal treating part 1 detecting memory of the memory accumulating part 3 by the memory collating part 2. The case where the shape L becomes below the reference value with which that it is impossible specific set up any value of similarity as a result of performing collation with L_m and calculating similarity with each shape, known shape L_1 , —, here is pointed out.

[0064]Although the shape L is difficult to specify as a result of search of the memory content of the memory accumulating part 3 at this time, the color C presupposes that it has specified.

Using the information on this color C, the hypothetical reasoning part 6 performs abduction reasoning, and builds up a hypothesis about an object (reasoning process (1): $L \rightarrow C$). And candidate S_1 of the object which has the color C in the pulsus cordis, —, S_k are enumerated (reasoning process (2): $S_1 \rightarrow C$, —, $S_k \rightarrow C$). These two reasoning process (1) – (2) is managed by the hypothetical reasoning 6. And even free needs to narrow down candidate S_1 of the object obtained in a reasoning process (2), —, S_k .

[0065]In the abduction in sign reasoning, since k syllogisms in a reasoning process (2) are processed on equal terms, it is difficult to narrow down even free, but. In the abduction in this example, it is considered as a candidate's object S_1 , —, the conditional probability $P(C|S_1)$ that connects S_k and the color C, —, the object which should presume S_g which gives the maximum in these probability values since $P(C|S_k)$ is used. As a result, it is presumed that the original object of the shape L is S_g (derivation hypothetical (3): $L \rightarrow S_g$). The hypothetical reasoning part 6 performs this candidate's narrowing down, and hypothetical derivation. And the information that an object is S_g is transmitted to the output deciding part 4 as it is via the memory collating part 2, and an action decision to S_g is made.

[0066]Thus, about derivation of the hypothesis by abduction reasoning, in the case of sign reasoning, narrowing down of two or more difficult candidate hypotheses becomes easy by using the pulsus cordis in this example, and it becomes possible to cancel the problem seen by sign reasoning.

[0067]Example 6. drawing 6 is an example in the case of having the condition-monitoring part 7 also as a formation part further in addition to the composition which showed the equipment configuration of this example and was shown in Example 1. They are the condition-monitoring part 7 and a treating part which supervises various internal parameters in the control system of the traveling vehicle which carries this example as an automatic navigation device. For example, as a result of running in the walking beat where this traveling vehicle was decided for a long time, it may fall into a fuel scarcity required for a run. In such a case, the condition-monitoring part 7 needs to detect the fuel scarcity of a traveling vehicle fuel tank, needs to interrupt work, and needs to make a traveling vehicle go to a fueling station. For this reason, the condition-monitoring part 7 transmits the command switched to the preferential manipulation routine for finding a fueling station to the external signal treating part 1 and the memory collating part 2.

[0068]For example, the map in a walking beat, the place of a fueling station, and scenery [in / further / all the turning points of the road in a zone] are memorized by the memory accumulating part 3 as picture images, If the geographic point in a zone can always check by comparing with the characteristic data processed by the external signal treating part 1 via the memory collating part 2, this traveling vehicle can arrive at a fueling station.

[0069]The external signal treating part 1 which received the manipulation-routine change command, and the memory collating part 2 switch to the mode management in a usually different emergency from the processing at the time. Although it waits, as for the difference between the processing at the time, and the processing in an emergency, for processing of the external signal treating part 1 to usually finish the memory collating part 2 as a part in the memory collating part 2 and collation with the memory content of the memory accumulating part 3 is performed, In an emergency, the memory collating part 2 collects only information required for the correspondence in an emergency among the memory contents of the memory accumulating part 3, and manages it as a memory content of the memory collating part 2, and processing is limited to feature extraction to the external signal treating part 1 based on the memory content.

[0070]For example, it is assumed that the rotating red lamp is a mark as a cue for finding a fueling station. It is required for the external signal treating part 1 that the memory collating part 2 should limit processing to the feature extraction for finding this mark preferentially. sometimes the external signal treating part 1 usually responds at necessity — shape, a color, and a motion — a texture, although it comes out and detailed feature extraction is performed, Processing is limited to the feature extraction about the motion for finding "what rotating" in

response to the demand from the memory collating part 2 in an emergency, and the feature extraction about the color for finding "red taro." At this time, processing is usually limited to "red taro" and a simple distinction called [or] other than this to having analyzed the detailed color for distinguishing various colors in an emergency, for example about a color by the processing at the time. The same may be said of the feature extraction about a motion.

[0071] Since the processing in an emergency in this example has received the command that it is a state of emergency from the condition-monitoring part 7 as mentioned above unlike the usual processing accompanied by a detailed collation process with the memory content of the memory accumulating part 3, It switches to the manipulation routine for arriving at a fueling station promptly through the collation process simplified comparatively.

[0072] Thus, when this information processor has the condition-monitoring part 7 which supervises various internal parameters of the state of the target system, it can respond to the accident inside a system and can respond to the state of emergency which a system faces flexibly by changing the usual information processing routine. As abnormality information, not only system inside information but system external information may be included.

[0073] Although the above example explained the case where the conversion table of the recognition object managed by the output deciding part 4 and an action plan was given beforehand, After this conversion table learns a situation with various information processors by this invention, it may be made to be gained, and necessary minimum knowledge is described to this conversion table, and whenever it is placed by various situations, correct this conversion table and it may be made to be extended it.

[0074] Although the example of vision information explained above the sensor information incorporated as an external signal in the example, other sensor information, for example, acoustic sense, sense of smell, taste information, etc. other than vision information may be used as sensor information.

[0075]

[Effect of the Invention] Since in-house-data expression which can be compared from an external signal can be created automatically according to the 1st composition or process of this invention, the object recognition which does not need human being's intervention, and decision-making based on it are possible. Since the collation system with high redundancy based on the appearing probability of attribution information is used, high object recognition or decision-making of pliability is more possible than the conventional sign reasoning method. By operation experience under the environment used, correction and extension of the stored data of a memory accumulating part are possible, and improvement of the reliability of recognizing ability and expansion can be performed in study.

[0076] According to the 2nd composition or process of this invention, by having a hypothetical reasoning part, also when the target presumption is impossible, presumption becomes possible by abduction reasoning. The huge calculation which is a problem in the abduction by the conventional signal reasoning is avoided by the use of a collation system based on the appearing probability of attribution information.

[0077] According to the 3rd composition or process of this invention, it becomes high-reliability-izing of the target recognizing ability, and what has an expansive more certain learning function by having an output evaluating part.

[0078] According to the 4th composition or process of this invention, the quick and suitable correspondence to abnormal circumstances is attained by having a condition-monitoring part.

.....
[Translation done.]

*** NOTICES ***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.*** shows the word which can not be translated.

3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1]It is a figure showing the information processor by one example of this invention.

[Drawing 2]It is a figure showing the image specimen procession in this example 1.

[Drawing 3]It is a figure showing the matching table managed by the output deciding part in this example 1.

[Drawing 4]It is a figure showing the information processor in the case of having an output evaluating part by one example of this invention.

[Drawing 5]It is a figure showing the information processor in the case of having a hypothetical reasoning part by one example of this invention.

[Drawing 6]It is a figure showing the information processor in the case of having a condition-monitoring part by one example of this invention.

[Drawing 7]It is a figure showing the composition of the inference system in a conventional system.

[Drawing 8]It is a figure showing the composition of the abduction inference system by the sign reasoning in a conventional system.

[Drawing 9]It is a figure showing the composition of the image recognition system in a conventional system.

[Drawing 10]It is a figure for the feature vector of the image recognition system in a conventional system to be shown.

[Drawing 11]It is a figure showing the representative picture image of recognition object operation of the image recognition system in a conventional system.

[Drawing 12]It is a figure showing the symbol corresponding to the representative picture image of the image recognition system in a conventional system.

[Drawing 13]It is a figure showing the example of the symbol trains corresponding to a certain operation processed with the image recognition system in a conventional system.

[Description of Notations]

1. External signal treating part
2. memory collating part
3. Memory accumulating part
4. output deciding part
5. Output evaluating part
6. hypothetical reasoning part
7. Condition-monitoring part

[Translation done.]

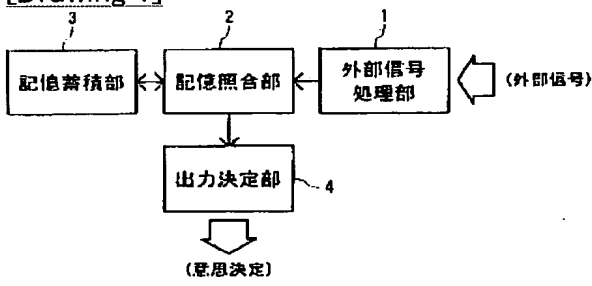
* NOTICES *

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.*** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DRAWINGS

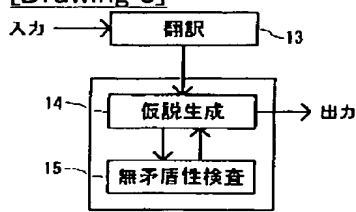
[Drawing 1]



[Drawing 2]

	S1	...	Si	...	Sn
L1	P(L1 S1)	...	P(L1 Si)	...	P(L1 Sn)
Lj	P(Lj S1)	...	P(Lj Si)	...	P(Lj Sn)
Lm	P(Lm S1)	...	P(Lm Si)	...	P(Lm Sn)

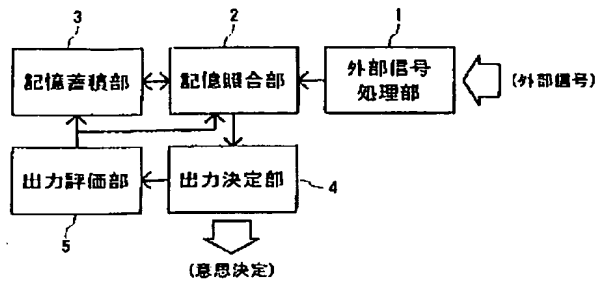
[Drawing 8]



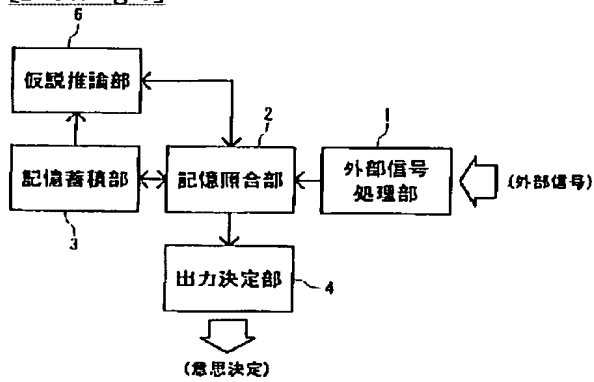
[Drawing 3]

対象	行動
人間	人間の声
ゴリラ	射撃音
チンパンジ	警笛
オランウータン	鈴の音

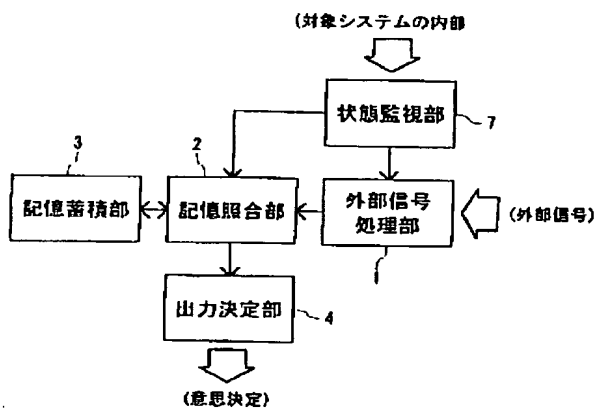
[Drawing 4]



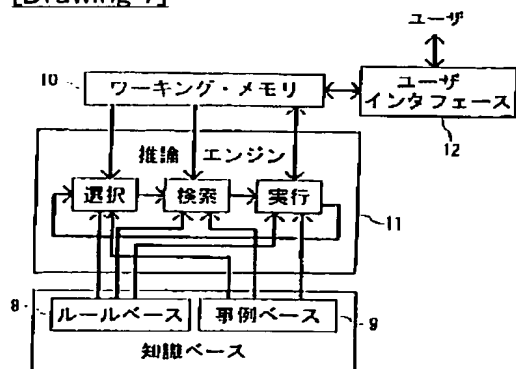
[Drawing 5]



[Drawing 6]



[Drawing 7]



[Drawing 9]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-305853

(43) 公開日 平成8年(1996)11月22日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 7/00 1/00		9061-5H	G 0 6 F 15/70 15/62	4 6 5 A 3 8 0

審査請求 未請求 請求項の数14 O L (全 12 頁)

(21) 出願番号 特願平7-105318

(22) 出願日 平成7年(1995)4月28日

(71) 出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72) 発明者 鹿毛 裕史

尼崎市塚口本町八丁目1番1号 三菱電機
株式会社中央研究所内

(72) 発明者 塩野 悟

尼崎市塚口本町八丁目1番1号 三菱電機
株式会社中央研究所内

(72) 発明者 山田 訓

尼崎市塚口本町八丁目1番1号 三菱電機
株式会社中央研究所内

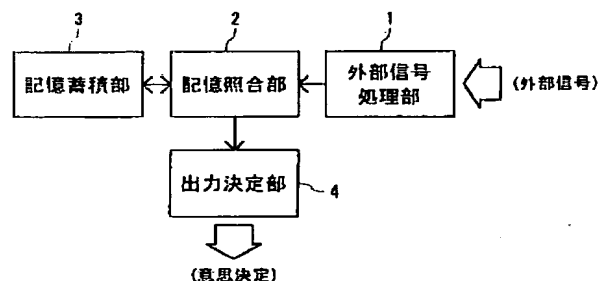
(74) 代理人 弁理士 高田 守 (外4名)

(54) 【発明の名称】 対象の認識およびそれにもとづく意思決定の方法およびその装置

(57) 【要約】

【目的】 人間の介入を必要とせず、外部情報にもとづいて対象を認識し、それに対する意思決定を行ない、環境での使用に従って記憶データの修正と拡張を柔軟に行う対象認識と意思決定の方法およびその装置を提供する。

【構成】 認識対象に関するセンサ情報から対象の属性情報を抽出し、照合可能な内部データ表現に変換する外部情報処理部と、複数の対象夫々の複数の属性情報の条件付出現確率を表現した内部データ表現を構成して記憶する記憶蓄積部と、外部情報処理部から新規に入力された対象の属性情報と記憶蓄積部に記憶された対象の属性情報との夫々の内部データ表現を照合する記憶照合部とを備えた。



【特許請求の範囲】

【請求項 1】 新規な認識対象に関するセンサ情報を外部信号として取込み、この外部信号から前記認識対象の属性情報を抽出し、この属性情報を照合可能な内部データ表現に変換する外部信号処理部と、複数の前記認識対象の夫々に関する複数の属性情報の条件付出現確率を表現した内部データ表現を構成して記憶する記憶蓄積部と、

前記外部信号処理部から入力された認識対象の属性情報の内部データ表現と前記記憶蓄積部の記憶データの内部データ表現とを照合する記憶照合部とを備えた情報処理装置を用い、

前記記憶蓄積部に記憶された認識対象のうち、前記外部信号処理部から新規に入力された認識対象と類似度の大きい内部データ表現を有する認識対象を前記新規に入力された認識対象であると推定する対象認識方法。

【請求項 2】 前記記憶蓄積部の内部データ表現として、対象毎の各属性情報の条件付出現確率からなるイメージ標本行列を用いる請求項 1 に記載の対象認識方法。

【請求項 3】 前記推定処理の経験に従い、前記記憶蓄積部の記憶データである条件付出現確率の修正、属性情報項目の追加または認識対象の追加を、前記推定処理の成功および失敗にもとづいて学習的に行なう請求項 1 または 2 に記載の対象認識方法。

【請求項 4】 前記情報処理装置に仮説推論部を備え、前記推定が不可能な新規に入力された認識対象に対して、前記推定に使用しなかった他の属性情報と前記記憶蓄積部の記憶データとの夫々の内部データ表現を用いたアブダクション推定を行って推定対象の候補を列挙し、前記新規入力対象と前記各候補対象との夫々の内部データ表現の類似度比較によって前記候補の絞り込みを行なう請求項 1 から 3 のいずれか 1 項に記載の対象認識方法。

【請求項 5】 新規な認識対象に関するセンサ情報を外部信号として取込み、この外部信号から前記認識対象の属性情報を抽出し、この属性情報を照合可能な内部データ表現に変換する外部信号処理部と、複数の前記認識対象の夫々に関する複数の属性情報の条件付出現確率を表現した内部データ表現を構成して記憶する記憶蓄積部と、前記外部信号処理部から入力された認識対象の属性情報の内部データ表現と前記記憶蓄積部の記憶データの内部データ表現とを照合する記憶照合部と、前記複数の認識対象と各認識対象に対してなすべき行動との対応表を有し、前記なすべき行動を命令として出力する出力決定部とを備えた情報処理装置を用い、前記記憶蓄積部に記憶された認識対象のうち、前記外部信号処理部から新規に入力された認識対象と類似度の大きい内部データ表現を有する認識対象を前記新規に入力された認識対象であると推定して認識し、

前記出力決定部の認識対象となすべき行動との対応表から前記認識された認識対象に対してなすべき行動を決定し、これを命令として出力する意思決定方法。

【請求項 6】 前記情報処理装置に出力評価部を備え、前記命令が実行された結果の認識対象の反応から前記命令または前記認識の妥当性を評価し、この評価に基づいて前記記憶蓄積部の記憶データである条件付出現確率の修正、属性項目の追加または認識対象の追加を学習的に行なう請求項 5 に記載の意思決定方法。

【請求項 7】 前記情報処理装置に状態監視部を備え、前記意思決定を行なう環境からの異変情報または前記センサからの異変情報に基づいて、前記推定処理および前記意思決定処理の内容をあらかじめ設定した緊急時用の内容に制限する請求項 5 または 6 に記載の意思決定方法。

【請求項 8】 新規な認識対象に関するセンサ情報を外部信号として取込み、この外部信号から前記認識対象の属性情報を抽出し、この属性情報を照合可能な内部データ表現に変換する外部信号処理部と、

複数の前記認識対象の夫々に関する複数の属性情報の条件付出現確率を表現した内部データ表現を構成して記憶する記憶蓄積部と、

前記外部信号処理部から入力された認識対象の属性情報の内部データ表現と前記記憶蓄積部の記憶データの内部データ表現とを照合する記憶照合部とを備え、

前記記憶蓄積部に記憶された認識対象のうち、前記外部信号処理部から新規に入力された認識対象と類似度の大きい内部データ表現を有する認識対象を前記新規に入力された認識対象であると推定する対象認識装置。

【請求項 9】 前記記憶蓄積部の内部データ表現として対象毎の各属性情報の条件付出現確率からなるイメージ標本行列を用いる請求項 8 に記載の対象認識装置。

【請求項 10】 前記推定処理の経験に従い、前記記憶蓄積部の記憶データである条件付出現確率の修正、属性情報項目の追加または認識対象の追加を、前記推定処理の成功および失敗にもとづいて学習的に行なう請求項 8 または 9 に記載の対象認識装置。

【請求項 11】 前記推定が不可能な新規に入力された認識対象に対して、前記推定に使用しなかった他の属性情報と前記記憶蓄積部の記憶データとの夫々の内部データ表現を用いたアブダクション推定を行って推定対象の候補を列挙する仮説推論部を備え、前記新規入力対象と前記各候補対象との夫々の内部データ表現の類似度比較によって前記候補の絞り込みを行なう請求項 8 から 10 のいずれか 1 項に記載の対象認識装置。

【請求項 12】 新規な認識対象に関するセンサ情報を外部信号として取込み、該外部信号から前記認識対象の属性情報を抽出し、この属性情報を照合可能な内部データ表現に変換する外部信号処理部と、複数の前記認識対象の夫々に関する複数の属性情報の条

件付出現確率を表現した内部データ表現を構成して記憶する記憶蓄積部と、
前記外部信号処理部から入力された認識対象の属性情報の内部データ表現と前記記憶蓄積部の記憶データの内部データ表現とを照合する記憶照合部と、
前記複数の認識対象と認識対象に対してなすべき行動との対応表を有し、前記なすべき行動を命令として出力する出力決定部とを備え、
前記記憶蓄積部に記憶された認識対象のうち、前記外部信号処理部から新規に入力された認識対象と類似度の大きい内部データ表現を有する認識対象を前記新規に入力された認識対象であると推定して認識し、
前記出力決定部の認識対象となすべき行動との対応表から前記認識された認識対象に対してなすべき行動を決定し、これを命令として出力する意思決定装置。

【請求項 13】 前記命令が実行された結果の認識対象の反応から前記命令または前記認識の妥当性を評価する出力評価部を備え、この評価に基づいて前記記憶蓄積部の記憶データである条件付出現確率の修正、属性項目の追加または認識対象の追加を学習的行なう請求項 12 に記載の意思決定装置。

【請求項 14】 前記意思決定を行なう環境からの異変情報または前記センサからの異変情報に基づいて前記推定処理および前記意思決定処理の内容をあらかじめ設定した緊急時用の内容に制限する状態監視部を備えた請求項 12 または 13 に記載の意思決定装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、外部信号から必要な情報を抽出し、その情報について記憶との照合、連想、推論を実行することによって状況に応じて対象を認識し、意思決定を行う情報処理の方法およびその装置に関するものである。

【0002】

【従来の技術】外界からの情報を用いて人間が行う意思決定を計算機に代行させるための研究がこれまで数多くなされてきた。それらは記号推論に基づく人工知能研究と、物体認識などを目標とする画像工学研究に大きく分けられる。外界から得た外部信号は画像信号とは限定されないが、用いられる割合が非常に大きく、ここでは画像信号に限定して述べる。前者の人工知能研究によって、エキスパート・システムなど多くの推論システムが構築され、意思決定支援システムとして実用化されてきた。後者の画像工学研究は、コンピュータ・ビジョンと呼ばれる研究分野を確立している。

【0003】推論システムの例として、人工知能学会誌 (1992), Vol. 7, No. 6, pp1087-1095 に掲載された安信千津子によるルールベースと事例ベースの両方を備え持つ知識ベース推論システムがある。従来のエキスパート・システムの多くは専門家の持

つ知識をルールベースに格納し、そのルールに基づいて推論を行うものであった。しかし一般にビジネス分野の審査業務などでは審査規制に定められていない事態も多く発生し、その場合過去の審査事例を参考にして判断する。このようにルールと併せて事例を活用することが判断に役立つ場合が多く、この従来例のシステムは状況に応じてルールもしくは事例を参考にして意思決定を支援する。

【0004】図 7 にこのシステムの構成を示す。8 は専門家の持つ知識を if-then ルールとして記憶しているルールベース、9 は過去の問題と解を事例として記憶している事例ベースである。10 はワーキング・メモリであり、現在の問題、解及び途中結果などのデータを記憶している。11 は推論エンジンであり、次の三つの機能を持つ。

〔1〕推論パラダイムの選択：ルールベース 8 もしくは事例ベース 9 に基づいた推論のいずれを選択するかについて、ワーキング・メモリ 10 の内容との照合処理によって決定し、現在直面している問題に対し適切な推論を行う。ルールベース 8 を構成するルールまたは事例ベース 9 を構成する事例と、ワーキング・メモリ 10 の内容とが適合するか否かを、0 から 1 の間の連続値を適合度として算出する。この照合結果に基づき、最も適合度の高いものがルールベース 8 のルールであるか事例ベース 9 の事例であるかによって、それぞれに基づいた推論を行う。両方のルールが同じ適合度の場合はルールベース 8 に基づいた推論を優先する。

〔2〕検索機能：現在の問題に対し適切な if-then ルールをルールベース 8 から検索するか、もしくは現在の問題に類似する事例を事例ベース 9 から検索する。検索は、上記の適合度計算と同様、各ルールまたは事例が役立つか否かについて、0 から 1 の間の連続値として計算する。

〔3〕実行機能：ルールベース 8 から検索された if-then ルールに従い実行するか、事例ベース 9 から検索した事例を現在の問題に適合するよう加工することにより、現在の問題の解を推論する。

【0005】12 はユーザインタフェースであり、ワーキング・メモリ 10 のデータの設定・変更、推論の起動の指示、推論結果の確認などの一連をユーザが行うための装置である。

【0006】上記の知識ベースシステムを始め、多くの推論システムでは演繹推論の枠内でしか推論が行えず、推論体系の拡張性に問題点があった。これを解決するために、記号推論の理論的研究から様々な推論形態が提唱されてきた。特にアブダクションという推論形態が近年注目を集めており、研究が進められている。

【0007】一般に人間は未知の状況に置かれた場合、過去の経験で得た知識に基づいて仮説を立て状況に応じた適切な判断を下す。アブダクションはこのような仮説

に基づく推論過程を基本としている。これにより従来の演繹や帰納推論とは異なり、経験的知識から新たに仮説を作り出し、その妥当性を試行錯誤により検証し新しい知識として取り込むことが可能である。アブダクション推論の形式は一般に次のように与えられる。

(1) $A \rightarrow B$ かつ (2) $C \rightarrow B$ ならば (3) $A \rightarrow C$

(1) は今直面している状況を示し、(2) は今までに経験して知っている知識の中の一つに対応し、(3) は(1) と(2) に基づいて導かれた仮説である。

【0008】(2) における C は B を導く候補として一般に数多く存在するので、結果として膨大な数の仮説が得られ、従ってその数を絞込むことが処理速度の点から重要となる。さらに得られた仮説が既に持っている論理体系と矛盾しないかどうかを検査する必要がある。こうした仮説の数の膨大さと無矛盾性の検査という二点を踏まえ、記号論理によるアブダクション推論を高速に処理することを目指したシステムとして、人工知能学会誌(1993) Vol.18, No.6, pp786-796 に掲載された井上克巳らによる例がある。

【0009】図8はアブダクション推論システムのモジュール構成である。13は述語論理で記述された論理式を高速計算のための記号表現に変換する翻訳モジュール、14は変換された記号表現から仮説を生成する仮説生成モジュール、15は生成された仮説が既に与えられている論理体系と矛盾しないかどうかを検査する無矛盾性検査モジュールである。ある仮説に対する矛盾の検査法には次の二通りがある。

(I) 論理式により記述された推論体系に基づき推論を行い、矛盾を検出する。

(II) 推論体系の中に矛盾を起こす仮説の集合を登録しておき、それとの照合により矛盾を検出する。この例では上記の(I)に基づいた無矛盾性検査をモジュール15によって処理し、そしてモジュール14とモジュール15の機能を論理型言語向きの高速な並列計算機上に実現し、並列に実行する計算機の台数を増やすことでアブダクション推論をスピードアップさせている。

【0010】上記の二つの従来例はいずれも記号推論システムであり、次のような問題が生じる。問題解決に必要な情報は全て記号表現として与えるので、外界情報から記号への変換と意味の付与は全て人間が行う必要がある。視覚情報など様々なセンサ情報が処理され、知識として表現する方法が与えられていなければ、いかなるシステムも常に人間の関与を必要とする。従って人間が行う意思決定を全て計算機に代行させることは不可能であり、解決可能な問題は極めて限定される。

【0011】このことから、センサ情報を処理し外界情報を内部表現に変換するシステムの構築が不可欠である。コンピュータ・ビジョンなどの画像工学研究では、静止画像または動画画像から物体の形状および動きを識別

し、さらに複雑な背景から視標を分離するためのモデルが提唱され、多くの画像認識システムが構築されてきた。これらのうち、動画画像処理により人間の行動認識を行う画像認識システムとして、特開平5-46583号公報に示された大和淳司らによる例がある。多くの画像認識システムでは、形状抽出などの低い認知レベルに留まっているのに対し、この従来技術では人間の行動認識といった高度な認知レベルを目標としている。

【0012】図9はこの従来例のアルゴリズムを表わす。まず、画像入力部16から行動中の人間を含む動画画像を捉え、画像用メモリ17に格納する。次に、特徴抽出部18により、動画画像から特徴ベクトルが得られる。ここで具体的な特徴ベクトルの計算方法の一例を説明する。

【0013】図10に示すように画像用メモリを $n \times m$ の画像数を持つ $N \times M$ のサブブロックに分割し、各々のサブブロックで画像の二値化を行う。次にこのサブブロック内の黒画素の占有率を求め、占有率を変数値とする $N \times M$ 次元の特徴ベクトルとする。この特徴ベクトルは特徴格納メモリ19に記録される。そして特徴ベクトルは量子化部20によってシンボルに変換され、シンボル格納メモリ21に記録される。この特徴ベクトルがシンボルに変換される方法を説明する。

【0014】図11は人間の動作認識のための代表的な画像例であり、図12は図11の各画像の特徴ベクトルに対応するシンボル(数字)である。時系列系に取り込まれた複数の画像データはそれぞれ特徴ベクトルに変換された後、特徴ベクトルに対応したシンボルの列に変換される。図13はある時系列的に取り込まれた図11とは別の画像データの列がシンボル列に変換された例を示す。量子化部20はあらかじめいくつかの代表点ベクトルを持っており、これらの一つ一つをシンボルと呼ぶ。つまり、量子化部20におけるシンボルへの変換とは、特徴ベクトルが代表点ベクトル群の中で最も距離の近い代表点ベクトルに対応するシンボルを選ぶということに相当する。

【0015】次に認識するカテゴリ数だけ用意された認識用状態遷移モデル格納メモリ25に格納されたモデルの各々から、特徴ベクトルが生成される確率を尤度算出部22によって算出する。求められた尤度が最大となるモデルが、認識結果として選択され認識結果用メモリ23に蓄積される。モデルパラメータ推定部24は、各カテゴリ毎に複数与えられた学習用データから得られたシンボルに対して、そのシンボルを発生するような状態遷移モデルのパラメータを推定し、認識用状態遷移モデル格納メモリ25に蓄積する。

【0016】この従来例では、認識対象の行動として四つの動作(左足を上げる、左手を上げる、右足を上げる、右手を上げる)を例に挙げ、認識率として平均約90%という効果が得られている。しかし識別可能な人間

の動作は数種類に限定され、また未知の動作を新たに一つの動作として登録する人間の動作は数種類に限定され、また未知の動作を新たに一つの動作として登録する手法は与えられておらず、認識できる動作の拡張性に問題があった。

【0017】

【発明が解決しようとする課題】従来の推論システムの例である安信らの方法は、if-thenルールに基づく演繹推論システムで、推論体系の拡張性に問題があった。その問題点を解決するためにアブダクション推論を導入した井上らの方法は、仮説の処理に関する膨大な量の計算を、専用並列計算機で高速に処理することによって対処している。しかし仮説の数の膨大さという問題は本質的に解決されていない。さらにこの二つの推論システムはいずれも記号推論システムであるため、問題解決に必要な情報は全て記号表現として人間が与えないといけないという問題点、すなわち各種センサで捉えられる外界情報の内部表現への変換、および意味の付与は全て人間の関与を必要とするという問題が残されている。

【0018】従来の画像認識システムの例である大和らの方法は、人間の関与しない認識の一手法を目指したものである。しかし識別可能な人間の動作は数種類に限定され、未知の動作への対応は考慮されておらず動作認識の柔軟性に問題があった。

【0019】本発明は上記のような従来の推論システムの画像認識システムがそれぞれ抱える問題点を解消するためになされたもので、与えられた状況に柔軟に応じて人間が関与せずに対象を認識し、その認識にもとづいて意思決定を行う情報処理方法およびその装置を提供することを目的とする。

【0020】

【課題を解決するための手段】上記の問題点を解決するために、人間がどのように外界を理解し意思決定をするかを詳細に検討する必要があった。そのため、人工知能や画像工学などの研究成果だけでなく、生理学や心理学及び論理学の分野の知見も考慮に入れた研究を行い、後述する心像という記憶表現（内部データ表現）を用いることを特徴とする情報処理の方法およびその装置を提案するに至った。

【0021】この発明に係わる情報処理装置は、例えば認識対象を含む視覚情報などのセンサ情報を外部信号として必要な情報を抽出した後、記憶部に格納した情報を使って心像を作りだし、その心像を使って認識対象を推定し、その結果に基づいて状況に応じた意思決定を行う。本情報処理装置の対象推定手段は外部信号処理部、記憶蓄積部、記憶照合部の少なくとも三つの部分から構成される。

【0022】第一に外部信号処理部について述べる。例えば外部信号が視覚情報である場合（音や臭いなどの他種の外部信号も同様に扱えるので、以下視覚情報に限定

して述べる）、それから内部データ表現を作る際に必要な情報を抽出する手段として、生理学的知見に基づいた視覚モデルを採用する。近年サル大脳皮質における視覚生理学研究から多くの知見が得られつつあり、そのメカニズムを模倣することにより人間に近い視覚システムを実現することが可能である。

【0023】サルの視覚関連野はいくつかの領域に分かれ、網膜情報の解析の局所性により階層構造をなしていることが知られている。一次視覚野では形状、動きおよび色などの網膜情報が局所的に解析され、他の領野で段階的に統合され、認識判断のための情報を作る。そこで本発明では、形状、動き、色およびテクスチャなどの属性を設け、CCDカメラなどで得た画像信号を処理して、属性に対応する情報を抽出して、システム内で照合可能な内部データ表現に変換する。

【0024】第二に、記憶蓄積部について述べる。記憶蓄積部における内部データ表現は、心理学研究により提案されている人間の記憶表現に近いものを用いている。ある三次元物体を認識対象とするとき、異なった視点からは異なった形状像が得られる。一般に認識対象は特定の視点から観察されることが多いので、対象の見え方に確率の差が生じる。その結果、対象が持つ形状像それぞれに条件付確率が伴っている。

【0025】本発明においては、一つの対象Sがもつ見え方の違いによる異なった形状像の集合と、観察者の過去の視覚体験を反映した対象Sが或る形状に見える条件付き確率との合わせて対象Sの心像と呼び、記憶蓄積部における内部データ表現として用いている。心像を構成する形状像などの属性情報は外部信号処理部で抽出された内部データ表現により構築される。

【0026】記憶蓄積部における、例えば形状に関する内部データ表現による記憶蓄積部の記憶内容の検索とは、心像を構成する形状像を検索することに相当する。動きなど他の属性についても同様である。記憶蓄積部の記憶内容である内部データ表現とその出現確率は固定したものではなく、経験により随時更新される。

【0027】第三に、記憶照合部について説明する。認識対象について外部信号処理部で作られ内部データ表現に変換された属性情報と記憶蓄積部の属性情報との類似度を使って、記憶蓄積部にある認識対象の候補を検索・抽出し、この記憶照合部にそれらの集合を作る。この集合を検索し、内部データ表現と最もよく一致するものを認識対象とする操作を行うことが、この記憶照合部の主な役割である。

【0028】認知心理学の研究によれば、人間の記憶は大きく短期記憶と長期記憶の二つに分けられる。短期記憶は最近に起こった出来事に関する記憶であり、長期記憶はほぼ固定的で一生を通じて保持される記憶である。これらの役割として、まず銘記される事象が短期記憶に取り込まれ、頻繁な想起によって強化された後、長

期記憶に固定されると考えられている。本発明の記憶照合部と記憶蓄積部は上述の短期記憶と長期記憶にそれぞれ対応する。

【0029】次に、本発明の意思決定手段に使用する出力決定部について述べる。出力決定部は、記憶照合部で認識対象と推定されたものに対して、なすべき行動を決定する。例えば認識対象が「火」であれば例えば「消火活動を行う」が対応づけられており、その他「近寄ってくる自動車」であれば「回避する」など、認識対象によって様々な行動パターンが対応づけられている。大脳生理学の研究によれば、人間の大脳皮質では感覚連合野で対象を認識し、その情報を前頭葉が受け、対象に応じてなすべき行動を決定する。本発明における出力決定部は、この前頭葉の機能に対応させている。

【0030】ここで出力決定部により決定される行動は、一般に既知の認識対象と対応している。しかし対象からの情報が不十分で認識が困難な場合や、または未知の対象に対しても仮説を立てて意思決定を行う必要がある。さらにその意思決定に基づく行動が妥当なものであったかどうかを評価する必要がある。このことから、本発明においては上記四種の主要構成部に加えて仮説推論部と出力評価部を設け、前者において内部データ表現が不十分で認識が不可能な場合にアブダクション推論を行って認識対象の候補を選び出し、後者においてそれに基づく行動の妥当性を評価する。さらに出力評価部を既知の認識対象に関する記憶蓄積部の記憶内容を変更したり、未知の対象を新たに記憶蓄積部に登録することにより、記憶蓄積部の持つ認識体系を拡張していく役割を持つ。

【0031】また対象とするシステムの内部状態に異変が生じ、それを回復するためには通常の情報処理ルーチンでは迅速に対処することが困難になる場合がある。これに対応する機能として、生物では喉の乾きや諸々の体内生理現象に直面した場合、視床下部と呼ばれる部位がその変化を感知し、迅速な行動に移るための意思決定に関わっていることが知られている。

【0032】大脳皮質を経由した判断は時間がかかるのに対し、視床下部の情報に基づく情動行動は大脳皮質を経由せず、緊急時の場合に即応的に対処する。本発明においては上記の主要四構成部に加えて、対象システム内部のパラメータや対象システム外部の状況の異変を監視する状態監視部を設けており、外部信号処理部や記憶照合部に緊急時処理ルーチンへの切り換えを命令し、システム内部状態の異変に即応的に対処することを可能にする。

【0033】

$$P(S_i | L_j) = P(L_j | S_i) P(S_i) / \{P(L_i | S_i) P(S_i) + \dots + P(L_j | S_n) P(S_n)\} \quad (1)$$

ここで、 $P(S_i)$ 、 \dots $P(S_n)$ はそれぞれ物体

$$P(S_i) = P(L_i | S_i) + \dots + P(L_n | S_i) \quad (2)$$

【作用】本発明による情報処理装置は、外部信号から内部データ表現に変換するための手法として生物の持つ情報処理機能を模倣するモデルを用いて構成したので、人間が介在することなく内部データ表現を獲得できる。その表現形式は、従来の記号表現とは異なる冗長性を持っているので、記号推論よりも柔軟な推論が可能であり、記号論理によるアブダクション推論で見られた爆発的に増加する計算量の問題を解決できる。

【0034】また未知の認識対象に対する判断を可能にするように構成したので、状況に応じて記憶体系に拡張及び修正を加え柔軟な対象認識とそれにもとづく意思決定を行う。さらに本情報処理装置が対象とするシステムが直面する可能性のある危険的な状況に即応的に対処するためのメカニズムを兼ね備えている。なお、本情報処理装置は単独の意思決定装置として使用できるだけでなく、意思決定を必要とする、一般のシステムに意思決定装置として組み込んで使用することができる。

【0035】

【実施例】

実施例 1. 図 1 は本実施例の構成を示す装置の例であり、外部信号処理部 1、記憶照合部 2、記憶蓄積部 3、出力決定部 4 の四部分からなる。本発明による情報処理装置では、外部信号処理部 1 により外部信号から必要な情報が抽出され、色や形状など属性毎に内部データ表現に変換される。そしてこの内部データ表現により、既に述べた記憶蓄積部 3 の記憶表現である心像を検索し、対象の推定を行う。

【0036】人間は生後様々な視覚体験を通じて、ある特定の物体の持つ様々な二次元形状像をその出現確率とともに記憶し、その後の物体の認識には形状像と過去の視覚体験に基づく出現確率を考慮して認識判断を行う。このようにいくつかの物体と様々な形状の対応関係を出現確率によって表現したものがイメージ標本行列と呼ばれる (Science(1992), Vol.257, pp1357-1363)。三次元物体 S_1, \dots, S_n が網膜像 L_1, \dots, L_n として映る場合、物体 S_i が網膜像 L_j として見える条件付き確率を $P(L_j | S_i)$ としたとき、イメージ標本行列は図 2 のように表される。

【0037】人間は物体 S_1, \dots, S_n に対する形状像 L_1, \dots, L_n の関連付けを図 2 のような条件付き確率の行列として獲得し、獲得後の視覚認知の際にはこの行列を行方向に参照することにより、特定の形状像 L_j から、元の物体が S_1, \dots, S_n のいずれであるかを推定する。イメージ標本行列を用いて、形状像 L_j が観測されたときの元の物体が S_i である確率 $P(S_i | L_j)$ は、ベイズ推定により次式で与えられる。

S_1, \dots, S_n の生起確率を示し、次式で表わされる。

各確率の値は過去の視覚体験を反映しており、状況依存的な値をとる。

【0038】以上の考え方を踏まえた上で、本実施例における三次元物体の推定方式を次のように考える。三次元物体 S_x は、見る方向等の条件の違いによって見え方の異なる二次元像の形状集合 $\{L_1, \dots, L_y\}$ を持っている。各々の形状像 L_j は物体 S_x と過去の視覚体験を反映した条件付き確率 $P(L_1 | S_x), \dots, P(L_y | S_x)$ によって関係づけられている。この形状集合を物体 S_x の心像と呼ぶ。観測された形状像 L_j から元の物体を推定するとき、心像の中に形状 L_j を含む物体をすべて候補とし、候補物体 $S_i (i = 1, \dots, n)$ 全てについて(1)式から推定確率 $P(S_i | L_j)$ を計算する。その確率が最大となる物体 S が L_j から推定される三次元物体である。

【0039】本実施例では、形状像に基づく対象物体推定と、それに対する行動決定について説明する。即ち、視覚情報の持つ様々な属性のうち形状による推定を優先させ、動きや色などの他の属性は対象推定の際には優先度が低いとする。

【0040】外部信号処理部1はCCDカメラと画像処理装置から構成される。前者で撮影された画像を後者を用いて処理し、認識対象物体の二次元像から形状 L を抽出する。

【0041】記憶蓄積部3は、認識対象とする物体 $S_i (i = 1, \dots, n)$ の外部信号処理部1を用いて得た形状像 $L_j (j = 1, \dots, m)$ と、その出現確率 $P(L_j | S_i)$ を内部データ表現として記憶している。即ち実際に本実施例を適用する環境下で、 S_i の様々な形状をあらかじめ外部信号処理部1を用いて L_j を得て、その出現確率も自動的に求めることによって記憶蓄積部3のメモリを構築する。この記憶蓄積部のメモリ内容は、以後の経験により形状像の追加もしくは出現確率の変更という形で変化させるような学習機構を組み込んでいる。出現確率の変更は、例えば対象の推定に成功した時に増加し、失敗した時に減少するようにする。失敗と成功の判断は、あらかじめ設定した類似度の基準値を超えるかどうかで行なう。

【0042】対象の推定に成功しない場合はその対象を未知の対象として登録し、その後の推定経験によって出現確率を構築する。この様にして推定する対象を学習的に拡張することができる。

【0043】記憶照合部2は外部信号処理部1で得た L に対して記憶蓄積部3を検索し、一定確率以上の L を形

$$P(\text{human} | L) = P(L | \text{human}) \cdot P(\text{human}) / \{P(L | \text{human}) \cdot P(\text{human}) + P(L | \text{chimp}) \cdot P(\text{chimp}) + P(L | \text{goril}) \cdot P(\text{goril})\}$$

他の動物に関する確率 $P(\text{chimp} | L)$ 、 $P(\text{goril} | L)$ についても同様である。この結果、例えば $P(\text{human} | L)$ の値が最大であれば元の対象は人間であると推

状として集合に含む物体 $S_i (i = 1, \dots, n)$ を全て列挙する。さらに記憶蓄積部3に記憶されている出現確率 $P(L | S_i)$ から、(1)式に従って元の物体の候補 S を推定する。

【0044】出力決定部4では、各物体に対して取る行動が対応表として管理されている。記憶照合部2で形状 L から推定された元の物体 S に対して起こすべき行動 A が出力決定部4により選択される。

【0045】これら構成部の具体的な動作を、霊長類を専門的に飼育している自然動物園で動物を監視し、危険箇所にいる動物に警告を発し、その場を立ち去らせるシステムを搭載した危険箇所管理システムの例で説明する。このシステムは観測される動物が人間であれば人間の声で注意を促し、ゴリラであれば射撃音を聞かせ、チンパンジーであれば警笛を鳴らすなどして追い払うためのシステムである。人間以外の各動物は対応する音刺激に対して、その場を立ち去らせるように条件付けされているとする。

【0046】外部信号処理部1はCCDカメラと形状抽出用画像情報処理装置を装備したもので、過去に見た様々な動物の形状が観測される条件付き確率を記録し、一方記憶蓄積部3にその結果を記憶させている。

【0047】まず外部信号処理部1が対象を捉え、対象の形状 L を抽出する。記憶照合部2は抽出された形状 L に基づき記憶蓄積部3の記憶を検索し、形状 L を心像の中に持つ対象について、対象 S の動物園内に存在する確率 $P(S)$ が設定した値以上であるものを候補として列挙する。

【0048】本実施例では、形状 L が既知の形状 L_1, \dots, L_m のいずれかと特定可能であるとする。形状 L は特定可能であるとは、既知の形状全てと照合を行い、各形状との類似度を計算した結果、いずれかの類似度の値が設定した基準値以上になり、その中で類似度の最大値を与える形状を形状 L から特定された形状と呼ぶ。

【0049】対象を解析した結果、人間、チンパンジー及びゴリラが候補として列挙されたとすると、記憶蓄積部3に記憶されている確率と(1)式によって元の対象を推定する。人間、チンパンジー及びゴリラが形状像 L としてそれぞれ抽出される確率 $P(L | \text{human})$ 、 $P(L | \text{chimp})$ 及び $P(L | \text{goril})$ と、人間、チンパンジーおよびゴリラがそれぞれ動物園内に存在する確率 $P(\text{human})$ 、 $P(\text{chimp})$ 及び $P(\text{goril})$ により、形状 L から元の対象が人間であると推定できる確率 $P(\text{human} | L)$ は次式により計算される。

定され、この情報は出力決定部4へ送信される。

【0050】出力決定部4では、記憶蓄積部3に記憶されて対象とそれに対して起こすべき行動との対応付け表

が管理されており、これに基づき行動決定がなされる。例えば出力決定部 4 は図 3 のような対応付け表を管理しているとする。この表から「人間の声による警告」という行動が選択される。

【0051】このように、二次元形状像として捉えられる対象を心像を用いて推定し、状況に応じて柔軟な意思決定を下すことが可能である。

【0052】実施例 2. 実施例 2 の構成は図 1 と同じで

$$P(S_i | L_j) = P(L_j | S_i) / \{P(L_j | S_1) + \dots + P(L_j | S_n)\} \quad (3)$$

【0053】本実施例による推定法は、(1) 式による推定よりも計算が速く、各対象 S_i ($i = 1, \dots, n$) がほぼ同様に出現する状況などで有効な推定法である。また各対象を公平に比較することになる。

【0054】実施例 3. 実施例 3 の構成は図 1 と同じである。実施例 1 では形状像 L によって認識対象 S を推定する例を示した。しかし CCD カメラなどで取り込まれた画像中にノイズが多い場合などは形状像による推定では信頼性が低いので、例えば色 C など形状以外の他の属性情報を併用して認識対象 S の推定に対する信頼性を向上させることが可能である。本実施例では、外部信号処理部 1 では形状と色の検出をする画像処理部を持ち、記憶蓄積部 3 は各認識対象について形状と色を出現確率を反映した条件付き確率とともに記憶しているとする。

【0055】実施例 1 と同じ適用例で説明する。外部信号処理部 1 により形状 L と色 C が抽出され、記憶照合部

「人間」の場合： $\{P(\text{human} | L) + P(\text{human} | C)\} / 2 = 0.45$

「チンパンジー」の場合： $\{P(\text{chimp} | L) + P(\text{chimp} | C)\} / 2 = 0.20$

20

「ゴリラ」の場合： $\{P(\text{goril} | L) + P(\text{goril} | C)\} / 2 = 0.15$

この計算結果によって、対象を「人間」とであると推定することが可能である。

【0057】この例では形状と色の情報を用いた推定の例を示した。しかし明るさが充分でないときは色による判断も信頼性が低い場合が多いので、更に動きの情報を用いて対象推定の信頼度を向上させることも可能である。この計算例では、各属性に関する確率の平均を相加平均として計算しているが、状況に応じて形状や色、動きに対する信頼度を評価し、重み付け平均として計算することも可能である。

【0058】このように、形状だけでなく動きや色などの属性の心像を用いることによって、記憶照合部 3 における対象推定の信頼性を大きく向上させることができる。

【0059】実施例 4. 図 4 は本実施例の装置構成を示し、実施例 1 で示した構成に加え、出力評価部 5 を構成部としても持つ場合の装置の例である。出力評価部 5 は、出力決定部 4 により下された意思決定が正しいかどうかを監視する機能を備えている。

【0060】実施例 1 同様の状況で観察された対象 S に

ある。実施例 1 では、形状像 L_j が観測されたときに元の物体が S_i であると推定できる確率 $P(S_i | L_j)$ は、(1) 式で与えられるベイズ推定で計算を行っていた。そして物体 S_1, \dots, S_n の生起確率 $S(S_i), \dots, P(S_n)$ は過去の視覚体験を反映した状況依存的な値であった。本実施例では、この各対象の出現確率がすべて同じ値であるとして計算を行っている。その結果 (1) 式は次式のように簡略化される。

(3)

2 によって記憶蓄積部 3 の記憶を検出し、形状による対象推定では人間、チンパンジー、ゴリラおよびテナガザルが、色による対象推定では人間、チンパンジー、ゴリラおよびニホンザルが候補対処として列挙されたとする。そこで共通する三つの候補である人間、チンパンジーおよびゴリラについて、推定確率を形状 L と色 C について計算し、次のような結果を得たとする。

形状 L : $P(\text{human} / L) = 0.20, P(\text{chimp} / L) = 0.30, P(\text{goril} / L) = 0.20$

色 C : $P(\text{human} / C) = 0.70, P(\text{chimp} / C) = 0.10, P(\text{goril} / C) = 0.10$

【0056】この場合、実施例 1 のように形状 L の推定確率だけを用いて認識対象 S を「チンパンジー」とであると判定してもよいが、形状による対象推定の信頼性が低い場合、色 C の推定確率も同時に用いて計算する。例えば形状 L および色 C の推定確率の平均を計算し、

対し、出力決定部 4 により「人間の声による警告」という行動が選択された場合を考える。この警告に対する対象 S の行動を出力評価部 5 が観察する。対象 S が人間であれば人間の声に即座に反応し、直ちにその場を立ち退くが、人間以外の動物であればそれぞれ個別の音刺激により立ち去るように条件付けされているので、人間の声で動物を立ち去らせることはできない。

【0061】出力評価部 5 によって対象 S の行動に変化がないと判断されると、出力評価部 5 は記憶蓄積部 3 で記憶されているイメージ標本行列における、対象 S が外部信号処理部 1 により形状像 L として観測される確率 $P(L/S)$ の値を減少させる。このようにして更新されたイメージ標本行列に基づき、もう一度対象 S を推定するように出力評価部 5 は記憶照合部 2 に対し処理を促す。イメージ標本行列を更新しても対象の推定に成功しない場合は、その対象を未知の対象として登録し、その後の推定経験によって出現確率を構築する。実施例 1 においてはイメージ標本行列の更新や新規対象の拡張を類似度基準による推定の成功や失敗によって起動していたが、本実施例では出力評価部 5 による対象の行動確認結

果に基づいて起動することができるので、より信頼性の高いイメージ標本行列の修正と拡張が可能となる。

【0062】このように出力評価部5により、出力決定部4により下された意思決定に基づく対象Sの行動が予測された行動であるかどうかを出力評価部5で判断することによって、記憶蓄積部3の記憶体系を修正し、与えられた状況に柔軟に対応して推定の信頼性を改善していく学習機能を実現することが可能となる。

【0063】実施例5. 図5は本実施例の装置構成を示し、実施例1で示した構成に加え、仮説推論部6を構成部として持つ場合の装置の例である。仮説推論部6の機能は、対象の認識が不可能の場合にそれが何であるかについて仮説を立てることである。本実施例は実施例3同様、外部信号処理部1で検出された形状と色に基づき対象の推定を行うが、実施例3と異なるのは、外部信号処理部1により抽出された形状Lが、記憶照合部2によって記憶蓄積部3の記憶を検出した結果、いずれの二次元形状にも特定できなかった場合を想定している点である。ここで形状Lが特定不可能とは、既知の形状L₁, ..., L_mとの照合を行い、各形状との類似度を計算した結果、いずれの類似度の値も設定した基準値以下になる場合を指す。

【0064】このとき記憶蓄積部3の記憶内容の検索の結果、形状Lは特定困難であるが、色Cは特定できたとする。この色Cの情報を用いて、仮説推論部6はアブダクション推論を行い、対象について仮説を立てる（推論過程（1）：L→C）。そして色Cを心像の中に持つ対象の候補S₁, ..., S_kを列挙する（推論過程（2）：S₁→C, ..., S_k→C）。この二つの推論過程（1）・（2）は仮説推論6で管理する。そして推論過程（2）で得られる対象の候補S₁, ..., S_kをただ一つに絞りこむ必要がある。

【0065】記号推論におけるアブダクションでは推論過程（2）におけるk個の推論式は対等に処理されるのでただ一つに絞りこむのは困難であるが、本実施例におけるアブダクションでは候補の対象S₁, ..., S_kと色Cを関係付ける条件付き確率P(C|S₁), ..., P(C|S_k)を用いるので、これらの確率値の中で最大値を与えるS_gを推定すべき対象とする。この結果、形状Lの元の対象はS_gであると推定される（導出仮説（3）：L→S_g）。この候補の絞り込みと仮説の導出は仮説推論部6が行う。そして対象がS_gであるという情報を記憶照合部2を経由してそのまま出力決定部4に送信してS_gに対する行動決定を下す。

【0066】このように、アブダクション推論による仮説の導出について、記号推論の場合では困難であった複数の候補仮説の絞り込みが、本実施例では心像を用いることにより容易になり、記号推論で見られた問題点を解消することが可能になる。

【0067】実施例6. 図6は本実施例の装置構成を示

し、実施例1で示した構成に加え、さらに状態監視部7を構成部としても持つ場合の例である。状態監視部7、本実施例を自動ナビゲーション装置として搭載している走行車の制御システム内の様々な内部パラメータを監視する処理部である。例えばこの走行車が決められた巡回区域内を長時間走行した結果、走行に必要な燃料不足に陥る場合がある。このような場合、状態監視部7は走行車燃料タンクの燃料不足を感知し、作業を中断して走行車を燃料補給所へ向かわせる必要がある。このため状態監視部7は外部信号処理部1と記憶照合部2に対して燃料補給所を見つけるための優先的な処理ルーチンに切り換える命令を送信する。

【0068】例えば、巡回区域内の地図と燃料補給所の場所、さらに区域内の道路の全ての分岐点における風景が画像イメージとして記憶蓄積部3に記憶されており、記憶照合部2を介して外部信号処理部1で処理された特徴データと照合することにより、区域内における地理的位置が常を確認できるとすれば、この走行車は燃料補給所に到達することができる。

【0069】処理ルーチン切り換え命令を受信した外部信号処理部1と記憶照合部2は、通常時の処理とは異なる緊急時の処理モードに切り換わる。記憶照合部2における通常時の処理と緊急時の処理との違いは、通常部では記憶照合部2は外部信号処理部1の処理が終わるのを待って記憶蓄積部3の記憶内容との照合を行うが、緊急時には記憶照合部2は記憶蓄積部3の記憶内容のうち、緊急時の対応に必要な情報だけをまとめて記憶照合部2の記憶内容として管理し、その記憶内容に基づいて外部信号処理部1に対して特徴抽出に処理を限定する。

【0070】例えば燃料補給所を見つけるための手掛りとして、回転する赤いランプが目印になっているとする。記憶照合部2は、この目印を優先的にを見つけるための特徴抽出に処理を限定するよう外部信号処理部1に要求する。外部信号処理部1は、通常時には必要に応じて形状、色、動きやテキストチャなどで詳細な特徴抽出を行っているが、緊急時には記憶照合部2からの要求を受けて、「回転するもの」を見つけるための動きに関する特徴抽出と、「赤いもの」を見つけるための色に関する特徴抽出に処理を限定する。この時、例えば色に関しては通常時の処理では様々な色を区別するための詳細な色の解析を行っていたのに対し、緊急時には「赤いもの」かそれ以外かという単純な区別に処理を限定する。動きに関する特徴抽出についても同様である。

【0071】本実施例における緊急時の処理は、上述のように記憶蓄積部3の記憶内容との詳細な照合過程を伴う通常の処理とは異なり、状態監視部7から緊急状態であるという命令を受けているので、比較的簡略化された照合過程を経て燃料補給所に迅速に到達するための処理ルーチンに切り換わる。

【0072】このように、本情報処理装置が対象とする

システムの状態の様々な内部パラメータを監視する状態監視部 7 を持つことにより、システム内部の異変に対応し、通常の情報処理ルーチンを変更することによりシステムの直面する緊急状態に柔軟に対応することができる。また異常情報としてはシステム内部情報に限らず、システム外部情報を含めてもよい。

【0073】以上の実施例では、出力決定部 4 で管理されている認識対象と行動計画の対応表が前もって与えられる場合について説明したが、この対応表は本発明による情報処理装置が様々な状況を学習した上で獲得されるようにしてもよく、また必要最小限の知識をこの対応表に記述しておき、様々な状況に置かれる度にこの対応表が修正・拡張されるようにしてもよい。

【0074】以上実施例では、外部信号として取り込むセンサ情報を視覚情報の例で説明したが、視覚情報以外の他のセンサ情報、例えば聴覚、嗅覚および味覚情報などをセンサ情報として用いてもよい。

【0075】

【発明の効果】本発明の第 1 の構成または工程によれば、外部信号からの照合可能な内部データ表現を自動的に作成できるので、人間の介入を必要としない対象認識とそれにもとづく意思決定が可能である。また、属性情報の出現確率にもとづく冗長性の高い照合方式を用いているので、従来の記号推論方式より柔軟性の高い対象認識または意思決定が可能である。また、使用される環境下での稼動経験によって記憶蓄積部の記憶データの修正および拡張が可能であり、認識機能の信頼性の改善と機能拡張を学習的行なうことができる。

【0076】本発明の第 2 の構成または工程によれば、仮説推論部を備えることにより、対象の推定が不可能であった場合にもアブダクション推論によって推定が可能となる。また、属性情報の出現確率にもとづく照合方式の利用により従来の信号推論によるアブダクションでの問題点である膨大な計算が避けられる。

【0077】本発明の第 3 の構成または工程によれば、出力評価部を備えることにより対象の認識機能の高信頼化と拡張の学習機能がより確実なものとなる。

【0078】本発明の第 4 の構成または工程によれば、状態監視部を備えることにより異常事態への迅速で適切な対応が可能となる。

【図面の簡単な説明】

【図 1】 本発明の一実施例による情報処理装置を示す図である。

【図 2】 本実施例 1 におけるイメージ標本行列を示す図である。

【図 3】 本実施例 1 における出力決定部で管理されている対応付け表を示す図である。

【図 4】 本発明の一実施例による出力評価部を持つ場合の情報処理装置を示す図である。

【図 5】 本発明の一実施例による仮説推論部を持つ場合の情報処理装置を示す図である。

【図 6】 本発明の一実施例による状態監視部を持つ場合の情報処理装置を示す図である。

【図 7】 従来方式における推論システムの構成を示す図である。

【図 8】 従来方式における記号推論によるアブダクション推論システムの構成を示す図である。

【図 9】 従来方式における画像認識システムの構成を示す図である。

【図 10】 従来方式における画像認識システムの特徴ベクトルを示すための図である。

【図 11】 従来方式における画像認識システムの認識対象動作の代表画像を示す図である。

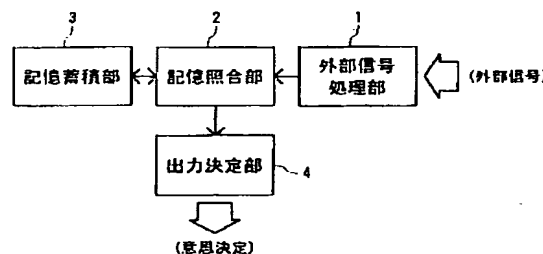
【図 12】 従来方式における画像認識システムの代表画像に対応するシンボルを示す図である。

【図 13】 従来方式における画像認識システムで処理されたある動作に対応するシンボル列の例を示す図である。

【符号の説明】

- | | |
|------------|----------|
| 1. 外部信号処理部 | 2. 記憶照合部 |
| 3. 記憶蓄積部 | 4. 出力決定部 |
| 5. 出力評価部 | 6. 仮説推論部 |
| 7. 状態監視部 | |

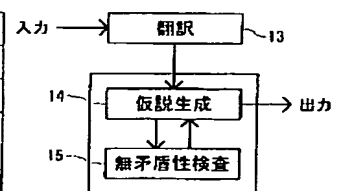
【図 1】



【図 2】

	S1	...	Si	...	Sn
L1	P(L1 S1)	...	P(L1 Si)	...	P(L1 Sn)
Lj	P(Lj S1)	...	P(Lj Si)	...	P(Lj Sn)
Lm	P(Lm S1)	...	P(Lm Si)	...	P(Lm Sn)

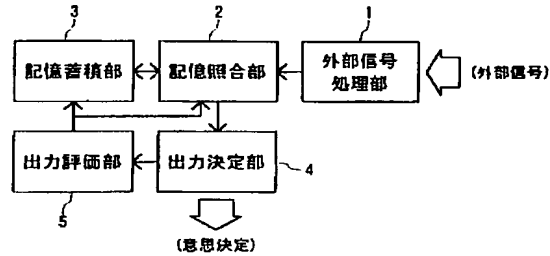
【図 8】



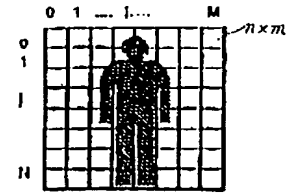
【図 3】

対象	行動
人間	人間の声
ゴリラ	射撃音
チンパンジ	警笛
オランウータン	鈴の音

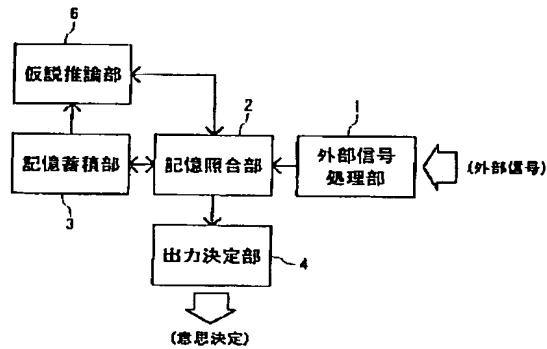
【図 4】



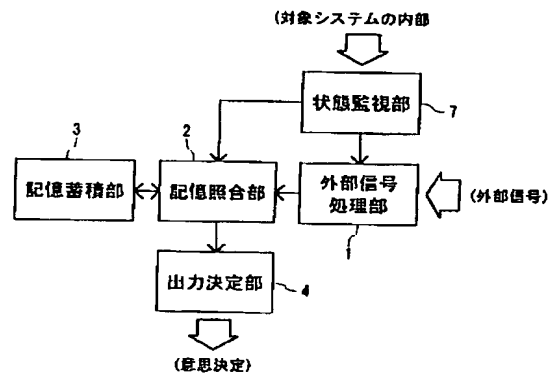
【図 10】



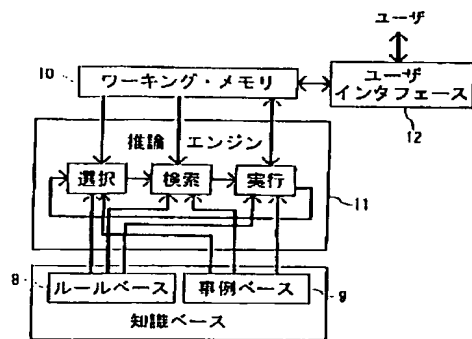
【図 5】



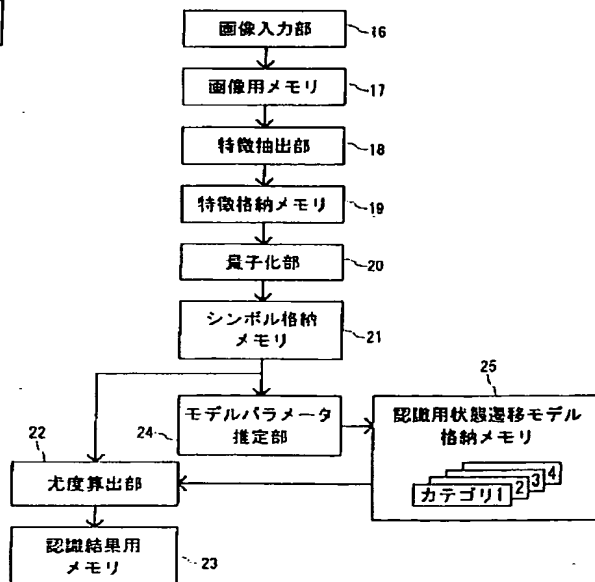
【図 6】



【図 7】



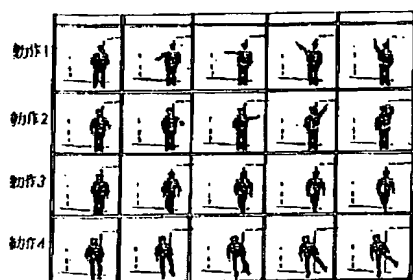
【図 9】



【図 12】

0	1	2	3	4
5	6	7	8	9
10	11	12	13	14
15	16	17	18	19

【图 1-1】



【図 13】

5 5 5 5 6 0 15 0 0 0 0 15 15 15 15 16 16 17 17 17 17 18 19 19 19 19 19 19 19
7 17 17 17 18 15 15 15 16 17 16 16 15 15 15 15 15 15 15 15 19 18 18 18 18 18 17